**Claims**

1. An isolated nucleic acid molecule selected from the group consisting of:

(a) nucleic acid molecules which hybridize under stringent conditions to a molecule consisting of a nucleotide sequence set forth as SEQ ID NO:1 and which code for a polypeptide having $C_\alpha$-formylglycine generating activity (FGE),

(b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code, and

(c) complements of (a) or (b).

2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises the nucleotide sequence set forth as SEQ ID NO:1.

3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule consists of the nucleotide sequence set forth as SEQ ID NO:3 or a fragment thereof.

4. An isolated nucleic acid molecule selected from the group consisting of

(a) unique fragments of a nucleotide sequence set forth as SEQ ID NO:1, and

(b) complements of (a).

5. The isolated nucleic acid molecule of claim 4, wherein the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.

6. The isolated nucleic acid molecule of claim 4, wherein the molecule encodes a polypeptide which is immunogenic.

7. An expression vector comprising the isolated nucleic acid molecule of claim 1, 2, 3, 4, 5, or 6, operably linked to a promoter.

8. An expression vector comprising the isolated nucleic acid molecule of claim 4 operably linked to a promoter.

9. A host cell transformed or transfected with the expression vector of claim 7.

10. A host cell transformed or transfected with the expression vector of claim 8.

11. An isolated polypeptide encoded by a nucleic acid molecule of claim 1, 2, 3, or 4, wherein the polypeptide, or fragment of the polypeptide, has $C_\alpha$-formylglycine generating activity.

12. The isolated polypeptide of claim 11, wherein the polypeptide is encoded by the nucleic acid molecule of claim 2.

13. The isolated polypeptide of claim 12, wherein the polypeptide comprises a polypeptide having the sequence of amino acids 1-374 of SEQ ID NO:2.

14. An isolated polypeptide encoded by a nucleic acid molecule of claim 1, 2, 3, or 4, wherein the polypeptide, or fragment of the polypeptide, is immunogenic.

15. The isolated polypeptide of claim 14, wherein the fragment of the polypeptide, or portion of the fragment, binds to a human antibody.

16. An isolated binding polypeptide which binds selectively a polypeptide encoded by an isolated nucleic acid molecule of claim 1, 2, 3, or 4.

17. The isolated binding polypeptide of claim 16, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids of SEQ ID NO:2.

18. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide is an antibody or an antibody fragment selected from the group consisting of a Fab fragment, a $F(ab)_2$ fragment or a fragment including a CDR3 region.

19. A family of isolated polypeptides having $C_\alpha$-formylglycine generating activity, each of said polypeptides comprising from amino terminus to carboxyl terminus:

(a) an amino-terminal subdomain 1;

(b) a subdomain 2 containing from 120 to 140 amino acids comprising at least 8 Tryptophans;

(c) a carboxy-terminal subdomain 3 containing from 35 to 45 amino acids;

wherein subdomain 2 has at least about 50% homology to subdomain 2 of a polypeptide selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78; and

wherein subdomain 3 has at least about 75% homology and a length approximately equal to subdomain 3 of a polypeptide selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78.

20. The polypeptides of claim 19, wherein subdomain 3 of each of said polypeptides has at least between about 80% and about 100% homology to subdomain 3 of a polypeptide selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78.

21. A method for determining the level of FGE expression in a subject, comprising measuring expression of FGE in a test sample from the subject to determine the level of FGE expression in the subject.

22. The method of claim 21, wherein the measured FGE expression in the test sample is compared to FGE expression in a control containing a known level of expression.

23. The method of claim 21, wherein the expression of FGE is FGE mRNA expression.

24. The method of claim 21, wherein the expression of FGE is FGE polypeptide expression.

25. The method of claim 21, wherein the test sample is tissue.

26. The method of claim 21, wherein the test sample is a biological fluid.

27. The method of claim 23, wherein FGE mRNA expression is measured using PCR.

28. The method of claim 23, wherein FGE mRNA expression is measured using Northern blotting.

29. The method of claim 24, wherein FGE polypeptide expression is measured using monoclonal antibodies to FGE.

30. The method of claim 24, wherein FGE polypeptide expression is measured using polyclonal antisera to FGE.

31. The method of claim 24, wherein expression of FGE is measured using $C_\alpha$-formylglycine generating activity.

32. A method for identifying an agent useful in modulating $C_\alpha$-formylglycine generating activity, comprising:

(a) contacting a molecule having $C_\alpha$-formylglycine generating activity with a candidate agent,

(b) measuring $C_\alpha$-formylglycine generating activity of the molecule, and

(c) comparing the measured $C_\alpha$-formylglycine generating activity of the molecule to a control to determine whether the candidate agent modulates $C_\alpha$-formylglycine generating activity of the molecule,

wherein the molecule is a nucleic acid molecule having a nucleotide sequence as the one set forth as SEQ ID NO:1, or an expression product thereof.

33. A method of diagnosing Multiple Sulfatase Deficiency in a subject, said method comprising:

(a) contacting a biological sample from a subject suspected of having Multiple Sulfatase Deficiency with an agent, said agent specifically binding to a molecule selected from the group consisting of: (i) a nucleic acid molecule having a nucleotide sequence as the one set forth as SEQ ID NO:1, (ii) an expression product of the nucleic acid molecule of (i), or (iii) a fragment of the expression product of (ii); and

b) measuring the amount of bound agent and determining therefrom if the expression of said nucleic acid molecule or of an expression product thereof is aberrant, aberrant expression being diagnostic of the Multiple Sulfatase Deficiency in the subject.

34. A method of diagnosing a condition characterized by aberrant expression of a nucleic acid molecule or an expression product thereof, said method comprising:

a) contacting a biological sample from a subject with an agent, wherein said agent specifically binds to said nucleic acid molecule, an expression product thereof, or a fragment of an expression product thereof; and

b) measuring the amount of bound agent and determining therefrom if the expression of said nucleic acid molecule or of an expression product thereof is aberrant, aberrant expression being diagnostic of the condition;

wherein the nucleic acid molecule has a nucleotide sequence as the one set forth as SEQ ID NO:1 and the condition is Multiple Sulfatase Deficiency.

35. A method for determining Multiple Sulfatase Deficiency in a subject characterized by aberrant expression of a nucleic acid molecule or an expression product thereof, comprising:

monitoring a sample from a patient for a parameter selected from the group consisting of

(i) a nucleic acid molecule having a nucleotide sequence as the one set forth as SEQ ID NO:1,

(ii) a polypeptide encoded by the nucleic acid molecule,

(iii) a peptide derived from the polypeptide, and

(iv) an antibody which selectively binds the polypeptide or peptide,

as a determination of Multiple Sulfatase Deficiency in the subject.

36. The method of claim 35, wherein the sample is a biological fluid or a tissue.

37. The method of claim 35, wherein the step of monitoring comprises contacting the sample with a detectable agent selected from the group consisting of

(a) an isolated nucleic acid molecule which selectively hybridizes under stringent conditions to the nucleic acid molecule of (i),

(b) an antibody which selectively binds the polypeptide of (ii), or the peptide of (iii), and

(c) a polypeptide or peptide which binds the antibody of (iv).

38. The method of claim 37, wherein the antibody, the polypeptide, the peptide or the nucleic acid is labeled with a radioactive label or an enzyme.

39. The method of claim 35, comprising assaying the sample for the peptide.

40. A kit, comprising a package containing:

an agent that selectively binds to the isolated nucleic acid of claim 1 or an expression product thereof, and

a control for comparing to a measured value of binding of said agent to said isolated nucleic acid of claim 1 or expression product thereof.

41. The kit of claim 40, wherein the control is a predetermined value for comparing to the measured value.

42. The kit of claim 40, wherein the control comprises an epitope of the expression product of the nucleic acid of claim 1.

43. The kit of claim 40, further comprising a second agent that selectively binds to a polypeptide selected from the group consisting of Iduronate 2-Sulfatase, Sulfamidase, N-Acetylgalactosamine 6-Sulfatase, N-Acetylglucosamine 6-Sulfatase, Arylsulfatase A, Arylsulfatase B, Arylsulfatase C, Arylsulfatase D, Arylsulfatase E, Arylsulfatase F, Arylsulfatase G, HSulf-1, HSulf-2, HSulf-3, HSulf-4, HSulf-5, and HSulf-6, or a peptide thereof, and

a control for comparing to a measured value of binding of said second agent to said polypeptide or peptide thereof.

44. A method for treating Multiple Sulfatase Deficiency in a subject, comprising:

administering to a subject in need of such treatment an agent that modulates $C_\alpha$-formylglycine generating activity, in an amount effective to treat Multiple Sulfatase Deficiency in the subject.

45. The method of claim 44, further comprising co-administering an agent selected from the group consisting of a nucleic acid molecule encoding Iduronate 2-Sulfatase, Sulfamidase, N-Acetylgalactosamine 6-Sulfatase, N-Acetylglucosamine 6-Sulfatase, Arylsulfatase A, Arylsulfatase B, Arylsulfatase C, Arylsulfatase D, Arylsulfatase E, Arylsulfatase F, Arylsulfatase G, HSulf-1, HSulf-2, HSulf-3, HSulf-4, HSulf-5, or HSulf-6, an expression product of the nucleic acid molecule, and a fragment of the expression product of the nucleic acid molecule.

46. The method of claim 44, wherein the agent that modulates $C_\alpha$-formylglycine generating activity is a nucleic acid molecule as claimed in Claims 1-8, or a nucleic acid having a sequence selected from the group consisting of SEQ ID NO: 1, 3, 4, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, and 80-87.

47. The method of claim 44, wherein the agent that modulates $C_\alpha$-formylglycine generating activity is a peptide as claimed in Claims 11-15, 19, 20, or a peptide having a sequence selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78.

48. The method of claim 44, wherein the agent that modulates $C_\alpha$-formylglycine generating activity is produced by a cell expressing an FGE nucleic acid molecule as claimed in Claims 1-8, or an FGE nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO: 1, 3, 4, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, and 80-87.

49. The method of claim 48, wherein the cell expressing an FGE nucleic acid molecule expresses an exogenous FGE nucleic acid molecule.

50. The method of claim 48, wherein the cell expressing an FGE nucleic acid molecule expresses an endogenous FGE nucleic acid molecule.

51. A method for increasing $C_\alpha$-formylglycine generating activity in a subject, comprising:

administering an isolated FGE nucleic acid molecule of the invention or an expression product thereof to a subject, in an amount effective to increase $C_\alpha$-formylglycine generating activity in the subject.

52. A method for treating a subject with Multiple Sulfatase Deficiency, comprising:

administering to a subject in need of such treatment an agent that modulates $C_\alpha$-formylglycine generating activity, in an amount effective to increase $C_\alpha$-formylglycine generating activity in the subject.

53. The method of claim 52, wherein the agent that modulates $C_\alpha$-formylglycine generating activity is a sense nucleic acid as claimed in Claims 1-8, or an FGE nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO: 1, 3, 4, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, and 80-87.

54. The method of claim 52, wherein the agent that modulates $C_\alpha$-formylglycine generating activity is an isolated polypeptide as claimed in Claims 11-15, 19, 20, or a peptide having a sequence selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78.

55. A method for increasing $C_\alpha$-formylglycine generating activity in a cell, comprising:

contacting the cell with an isolated nucleic acid molecule of claim 1 or an expression product thereof, in an amount effective to increase $C_\alpha$-formylglycine generating activity in the cell.

56. A pharmaceutical composition, comprising:

an agent comprising an isolated nucleic acid molecule as claimed in any one of Claims 1-8, an FGE nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO: 1, 3, 4, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, and 80-87, or an expression product thereof, in a pharmaceutically effective amount to treat Multiple Sulfatase Deficiency, and

a pharmaceutically acceptable carrier.

57. A method for identifying a candidate agent useful in the treatment of Multiple Sulfatase Deficiency, comprising:

determining expression of a set of nucleic acid molecules in a cell or tissue under conditions which, in the absence of a candidate agent, permit a first amount of expression of the set of nucleic acid molecules, wherein the set of nucleic acid molecules comprises at least one nucleic acid molecule selected from the group consisting of

(a) nucleic acid molecules which hybridize under stringent conditions to a molecule consisting of a nucleotide sequence set forth as SEQ ID NO:1 and which code for a polypeptide having $C_\alpha$-formylglycine generating activity (FGE),

(b) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code,

(c) a nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO: 1, 3, 4, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, and 80-87, and

(d) complements of (a) or (b) or (c),

contacting the cell or tissue with the candidate agent, and

detecting a test amount of expression of the set of nucleic acid molecules, wherein an increase in the test amount of expression in the presence of the candidate agent relative to the first amount of expression indicates that the candidate agent is useful in the treatment of the Multiple Sulfatase Deficiency.

58. A solid-phase nucleic acid molecule array consisting essentially of a set of nucleic acid molecules, expression products thereof, or fragments thereof, each nucleic acid molecule encoding for a polypeptide selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78, Iduronate 2-Sulfatase, Sulfamidase, N-Acetylgalactosamine 6-Sulfatase, N-Acetylglucosamine 6-Sulfatase, Arylsulfatase A, Arylsulfatase B, Arylsulfatase C, Arylsulfatase D, Arylsulfatase E, Arylsulfatase F, Arylsulfatase G, HSulf-1, HSulf-2, HSulf-3, HSulf-4, HSulf-5, and HSulf-6, fixed to a solid substrate.

59. The solid-phase nucleic acid molecule array of claim 58, further comprising at least one control nucleic acid molecule.

60. The solid-phase nucleic acid molecule array of claim 58, wherein the set of nucleic acid molecules comprises at least one nucleic acid molecule encoding for a polypeptide selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78, Iduronate 2-Sulfatase, Sulfamidase, N-Acetylgalactosamine 6-Sulfatase, N-Acetylglucosamine 6-Sulfatase, Arylsulfatase A, Arylsulfatase B, Arylsulfatase C, Arylsulfatase D, Arylsulfatase E, Arylsulfatase F, Arylsulfatase G, HSulf-1, HSulf-2, HSulf-3, HSulf-4, HSulf-5, and HSulf-6.

61. The solid-phase nucleic acid molecule array of claim 58, wherein the set of nucleic acid molecules comprises at least two nucleic acid molecules, each nucleic acid molecule encoding for a polypeptide selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78, Iduronate 2-Sulfatase, Sulfamidase, N-Acetylgalactosamine 6-Sulfatase, N-Acetylglucosamine 6-Sulfatase, Arylsulfatase A, Arylsulfatase B, Arylsulfatase C, Arylsulfatase D, Arylsulfatase E, Arylsulfatase F, Arylsulfatase G, HSulf-1, HSulf-2, HSulf-3, HSulf-4, HSulf-5, and HSulf-6.

62. A method for increasing sulfatase activity in a cell, comprising:

contacting a cell expressing a sulfatase with an isolated nucleic acid molecule as claimed in Claims 1-8, or a nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO: 1, 3, 4, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, and 80-87, or an expression product thereof, in an amount effective to increase sulfatase activity in the cell.

63. The method of claim 62, wherein the cell expresses endogenous sulfatase.

64. The method of claim 62, wherein the cell expresses exogenous sulfatase.

65. The method of claim 63, wherein the endogenous sulfatase is activated.

66. The method according to any one of claims 62-66, wherein the sulfatase is selected from the group consisting of Iduronate 2-Sulfatase, Sulfamidase, N-Acetylgalactosamine 6-Sulfatase, N-Acetylglucosamine 6-Sulfatase, Arylsulfatase A, Arylsulfatase B, Arylsulfatase C, Arylsulfatase D, Arylsulfatase E, Arylsulfatase F, Arylsulfatase G, HSulf-1, HSulf-2, HSulf-3, HSulf-4, HSulf-5, and HSulf-6.

67. The method of claim 62, wherein the cell is a mammalian cell.

68. A pharmaceutical composition, comprising:

a sulfatase that is produced by cell, in a pharmaceutically effective amount to treat a sulfatase deficiency, and

a pharmaceutically acceptable carrier,

wherein said cell has been contacted with an agent comprising an isolated nucleic acid molecule as claimed in Claims 1-8, or a nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO: 1, 3, 4, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, and 80-87, or an expression product thereof.

69. An isolated variant allele of a human FGE gene, which encodes a variant FGE polypeptide, comprising:

an amino acid sequence comprising at least one variation in SEQ ID NO:2, wherein the at least one variation comprises: Met1Arg; Met1Val; Ser155Pro; Cys218Tyr; Ala279Val;

Arg327Stop; Cys336Arg; Arg345Cys; Arg349Trp; Arg349Trp; Arg349Gln; Ser359Stop; or a combination thereof.

70. An isolated variant human FGE polypeptide, comprising:

an amino acid sequence comprising at least one variation in SEQ ID NO:2, wherein the at least one variation comprises: Met1Arg; Met1Val; Ser155Pro; Cys218Tyr; Ala279Val; Arg327Stop; Cys336Arg; Arg345Cys; Arg349Trp; Arg349Trp; Arg349Gln; Ser359Stop; or a combination thereof.

71. An antibody having the variant human FGE polypeptide of claim 69 as an immunogen.

72. The antibody of claim 71, which is a polyclonal antibody.

73. The antibody of claim 71, which is a monoclonal antibody.

74. The antibody of claim 71, which is a chimeric antibody.

75. The antibody of claim 71, detectably labeled.

76. The antibody of claim 75, wherein said detectable label comprises a radioactive element, a chemical which fluoresces, or an enzyme.

77. A sulfatase-producing cell wherein the ratio of active sulfatase to total sulfatase produced by the cell is increased, the cell comprising:

(i) a sulfatase with an increased expression, and

(ii) a Formylglycine Generating Enzyme with an increased expression,

wherein the ratio of active sulfatase to total sulfatase produced by the cell is increased by at least 5% over the ratio of active sulfatase to total sulfatase produced by the cell in the absence of the Formylglycine Generating Enzyme.

78. The method of claim 77, wherein the ratio of active sulfatase to total sulfatase produced by the cell is increased by at least 10% over the ratio of active sulfatase to total sulfatase produced by the cell in the absence of the Formylglycine Generating Enzyme.

79. The method of claim 77, wherein the ratio of active sulfatase to total sulfatase produced by the cell is increased by at least 20% over the ratio of active sulfatase to total sulfatase produced by the cell in the absence of the Formylglycine Generating Enzyme.

80. The method of claim 77, wherein the ratio of active sulfatase to total sulfatase produced by the cell is increased by at least 50% over the ratio of active sulfatase to total sulfatase produced by the cell in the absence of the Formylglycine Generating Enzyme.

81. The method of claim 77, wherein the ratio of active sulfatase to total sulfatase produced by the cell is increased by at least 100% over the ratio of active sulfatase to total sulfatase produced by the cell in the absence of the Formylglycine Generating Enzyme.

82. In a method for treating a sulfatase deficiency with a sulfatase by administering to a subject in need of such treatment a sulfatase to treat the sulfatase deficiency, the improvement comprising administering to the subject a sulfatase contacted with a Formylglycine Generating Enzyme in an amount effective to increase the specific activity of the sulfatase.

83. The method of claim 82, wherein the sulfatase is selected from the group consisting of Iduronate 2-Sulfatase, Sulfamidase, N-Acetylgalactosamine 6-Sulfatase, N-Acetylglucosamine 6-Sulfatase, Arylsulfatase A, Arylsulfatase B, Arylsulfatase C, Arylsulfatase D, Arylsulfatase E, Arylsulfatase F, Arylsulfatase G, HSulf-1, HSulf-2, HSulf-3, HSulf-4, HSulf-5, and HSulf-6.

84. The method of claim 82, wherein the Formylglycine Generating Enzyme is encoded by a nucleic acid molecule as claimed in Claims 1-8, or a nucleic acid having a sequence selected from the group consisting of SEQ ID NO: 1, 3, 4, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, and 80-87.

85. The method of claim 82, wherein the Formylglycine Generating Enzyme is a peptide as claimed in Claims 11-15, 19, 20, or a peptide having a sequence selected from the group consisting of SEQ ID NO. 2, 5, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, and 78.